NATIONAL ACADEMIES Sciences Engineering Medicine

SECURING AI SYSTEMS: New Challenges and Priorities

NSF Grant: **#2538196** -- *EAGER: Toward a Research Agenda for Securing Artificial Intelligence*

Context and Motivation:

Artificial Intelligence (AI) is advancing rapidly, presenting complex challenges for ensuring system security. Although AI systems are being widely adopted across sectors, comprehensive frameworks for safeguarding these technologies remain underdeveloped. Addressing this gap requires bringing together experts from across the research community to systematically assess current and emerging threats to AI systems, as well as the risks introduced by their deployment and application.

The project comprises two components:

1. A national convening of leading experts for a conversations toward characterizing challenges and research priorities, and

2. Development of a National Academies *issue paper* that outlines a forward-looking AI security research agenda for the community.

The community conversation and subsequent *issue paper* will serve as a foundational reference to help coordinate and guide efforts toward meaningful progress in addressing AI security challenges.

Key Outcomes To Date:

The National Academies of Sciences, Engineering, and Medicine convened the community on April 20-21, 2026 in Washington DC.

The meeting brought together ~ 70 experts in person, spanning industry, government, and academia. Over 200 participants joined the conversation online.

Meeting topics included definitions and frameworks for AI security, threat modeling, risk assessment, agentic AI security, and infrastructure supporting research and development.

Next Steps:

The National Academies are gathering input to develop an *issue paper* outlining new challenges and research priorities for securing AI systems. The issue paper is anticipated by the Fall of 2026.

Share your insights:

We welcome contributions to the upcoming *issue paper* from the CICI PI community. Those interested in contributing can send a note to Tho Nguyen at thonguyen@nas.edu.